

Fast adaptive flat-histogram ensemble to enhance the sampling in large systems

XU Shun¹, ZHOU Xin^{1*}, JIANG Yi² & WANG YanTing³

¹ School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China;

² Department of Mathematics and Statistics, Georgia State University, Atlanta 30303, USA;

³ State Key Laboratory of Theoretical Physics, Institute of Theoretical Physics,
Chinese Academy of Sciences, Beijing 100190, China

Received February 27, 2015; accepted April 29, 2015

An efficient novel algorithm was developed to estimate the Density of States (DOS) for large systems by calculating the ensemble means of an extensive physical variable, such as the potential energy, U , in generalized canonical ensembles to interpolate the interior reverse temperature curve $\beta_s(U) \equiv \frac{\partial S(U)}{\partial U}$, where $S(U)$ is the logarithm of the DOS. This curve is computed with different accuracies in different energy regions to capture the dependence of the reverse temperature on U without setting prior grid in the U space. By combining with a U -compression transformation, we decrease the computational complexity from $O(N^{3/2})$ in the normal Wang Landau type method to $O(N^{1/2})$ in the current algorithm, as the degrees of freedom of system N . The efficiency of the algorithm is demonstrated by applying to Lennard Jones fluids with various N , along with its ability to find different macroscopic states, including metastable states.

molecular dynamics simulations, enhanced sampling, density of states, generalized canonical ensemble, flat-histogram ensemble

PACS number(s): 05.10.-a, 02.70.Ns, 64.70.D-

Citation: Xu S, Zhou X, Jiang Y, et al. Fast adaptive flat-histogram ensemble to enhance the sampling in large systems. *Sci China-Phys Mech Astron*, 2015, 58: 590501, doi: 10.1007/s11433-015-5690-7

1 Introduction

An essential problem for molecular simulations in complex systems with a rugged potential energy surface, such as proteins [1,2] and glasses [3] or first-order phase-coexistence systems, is how to adequately sample the conformational space, since the existing high free-energy barriers prevent the regular Monte Carlo (MC) and Molecular Dynamics (MD) simulations to visit all the important regions sufficiently. In attempting to overcome this difficulty, many enhanced sampling methods [4–18] have been developed;

most of them tried to modify the potential energy surface to make the distribution function in a preselected collective variable space be almost flat. Although these methods have achieved great accomplishments, more attempts are still needed to further improve the sampling efficiency, especially in larger systems.

The Density of States (DOS) of a system is defined as $\Omega(U) \equiv \int dr \delta(U - U(r))$, where $U(r)$ is the potential energy of the system (or more generally, any conformational function). Here, r is a simple denotation of high-dimensional conformations, such as all-atomic position vectors, and the integral about r is over the whole conformational space. Usually, we define the microcanonical entropy $S(U) \equiv \ln$

*Corresponding author (email: xzhou@ucas.ac.cn)

$\Omega(U)$ and the interior reverse temperature $\beta_s(U) \equiv \frac{\partial S(U)}{\partial U}$

to describe the inner property of the system. Then, we have the flat-histogram ensemble, where the conformational distribution $P_{fh}(r) \propto e^{-S(U(r))}$. The standard simulation techniques, such as MC or MD simulations, can generate the desired distribution if a new potential $V_{new}(r) = T_0 S(U)$ is applied, where T_0 is the temperature of thermostat, which could be any value. Thus the histogram in the U space $P(U) \propto e^{-S(U)} \int dr P(r) \delta(U - U(r))$ is flat. Since the flat-histogram ensemble can visit a much wider energy range than the usual Canonical Ensemble (CE) where the conformational distribution $P_{ce}(r) \propto e^{-\beta U(r)}$, it has been widely used to overcome free-energy barriers and thus enhances the sampling for complex systems [4–6,9,10,13]. Here, $\beta = 1/k_B T$, with k_B Boltzmann constant and T the temperature. In the flat-histogram ensemble, $S(U)$ is an input quantity and the target of simulations at the same time, so $S(U)$ should be solved iteratively. The convergence efficiency and the accuracy of the iteration are essential for different flat-histogram methods.

Among them, the Wang Landau (WL) algorithm [5] lifts the value of S in the current-visited U range at each simulation step until the visiting histogram in U space is almost flat. The WL algorithm is easy to implement and has been widely applied to various kinds of systems. However, the number of required simulation steps is at least in the order of $O(N^2)$ [8] with the system size N . The recently developed Statistical-Temperature Molecular Dynamics (STMD) method [10] updates the intensive variable $T_s(U) \equiv \beta_s^{-1}(U)$ instead of the extensive variable $S(U)$ in the original WL, thus improving the efficiency of convergence to the order of $O(N^{3/2})$. A short point of these WL-like methods is that they do not satisfy the detailed balance condition thus with less numerical stability in applications. Alternatively, the multi-micro canonical method (MMCM) [7] calculates the microcanonical temperature $T_m(E)$ and relates the $T_m(E)$ to the interior reverse temperature $\beta_s(E)$ based on the thermodynamic equivalence of differently defined temperatures and can converge within $O(N^{3/2})$ simulation steps. Different from the original WL algorithm and the STMD, the MMCM satisfies the detailed balance condition, then has some benefits in numerical stability. In the WL-like algorithms and the MMCM, it is necessary to cut the extensive variable U into homogeneous bins with a presumed bin size. We collect sufficient simulation data in each bin of U to estimate the value of $\beta_s(U)$ at the centers of each U bin, that is, preset a uniform grid $\{U_i\}$, $i=1, \dots, M$ and calculate $\beta_s(U)$ on the grid, then interpolate the whole curve $\beta_s(U)$ (or $T_s(U)$). Since the varying rates of $\beta_s(U)$ at different U regions may be very different, the presetting of a suitable U grid for achieving both the accuracy and the efficiency is not trivial. We

should use a sparse $\{U_i\}$ grid in the slow-varied regions of $\beta_s(U)$, but a more dense grid in the fast-varied regions of $\beta_s(U)$.

In this paper, we present the Fast Adaptive Flat-histogram Ensemble (FAFE) method to efficiently estimate $\beta_s(U)$. The FAFE calculates the ensemble means and fluctuations of U in a series of generalized ensembles to adaptively form a U grid and to get the values of $\beta_s(U)$ on the grid, $\{(\beta_i, U_i)\}$. The $\{U_i\}$ grid is non-uniform and the grid bin widths depend on both U and N automatically: dense $\{U_i\}$ points are applied in the $\beta_s(U)$ fast-changed regions but sparse $\{U_i\}$ in the U regions where β_s varies slowly as U . In the U regions where $\beta_s(U)$ was already well estimated, we apply a compression transformation to further depress the visit time in that regions so that more simulation costs are focus on the others. By doing so, in the FAFE, the estimate of $\beta_s(U)$ can be converged within the order of $N^{1/2}$ simulation steps. Also, the FAFE satisfies the detailed balance condition, and it can visit and identify coexistent phases, including metastable phases of systems.

2 Theory and method

When an estimate of the entropy $\tilde{S}(U)$ is obtained, the CE means of any conformational function $Q(r)$ can be easily calculated by

$$\langle Q \rangle_T = \frac{\sum_k Q_k w_k}{\sum_k w_k}, \quad (1)$$

where $Q_k = Q(r_k)$ is the value of the $Q(r)$ at the sampled conformation r_k under the flat-histogram simulation, and the weight factor is $w_k = e^{\tilde{S}(U(r_k)) - U(r_k)/k_B T}$. It is worth mentioning that the weighted average is efficient only while the effective size of the weighted sample is large enough, which usually requires the generated sample to cover the target sample, that is, the visited U range of the flat-histogram ensemble sample covers that of the CE. We can approximately estimate the statistical error of the weighted average by the effective size of the weighted sample:

$$B = \frac{(\sum_k w_k)^2}{\sum_k w_k^2}, \quad (2)$$

which decreases with increasing fluctuation of normalized weights. While the weights are constant, the effective size of the sample, B , is the realistic size of the sample, as per our expectation. Only while all significant regions of the target distribution are adequately sampled, the generated sample can be weighted efficiently to reproduce the target one, that is, the effective size B of the weighted sample is much larger than unity.

For any $f(U)$ ensemble, that is, $P_f(r) \propto e^{-f(U(r))}$, the histogram $H_f(U) \propto e^{S(U)-f(U)}$ could be approximately written as:

$$H_f(U) \propto \exp\left[-\frac{1}{2\sigma_f^2}(U - E_f)^2\right], \quad (3)$$

by Taylor expansion, if N is very large, and $U \propto N$ is an extensive variable. The maximal point of $H_f(U)$ is approximately equal to the ensemble average of U under the $f(U)$ ensemble, $E_f \approx \langle U \rangle_f$. Thus, we have

$$\begin{aligned} \beta_s(E_f) &= f'(E_f), \\ \beta_s'(E_f) &= -\sigma_f^{-2} + f''(E_f), \end{aligned} \quad (4)$$

where the single prime and double prime mean the first derivative and the second derivative, respectively. σ_f is the fluctuation of U in the $f(U)$ ensemble. We calculate $E_f = \langle U \rangle_f$ and $\sigma_f^2 = \langle (U - E_f)^2 \rangle_f$ in a series of different $f(U)$

ensembles to generate points on the curve $\beta_s(U)$. The whole curve $\beta_s(U)$ can be interpolated from these points, and $\tilde{S}(U)$ can be estimated by integrating $\beta_s(U)$.

Different $f(U)$ ensembles, such as the CE $f(U) = \beta U$ and the microcanonical ensemble

$$f_{mce}(U) = -\frac{l-2}{2} \ln[(E-U)\Theta(E-U)] \quad (5)$$

can be used to construct $\beta_s(U)$. Here, l is number of degrees of freedom, the Heaviside step function $\Theta(x) = 1$ while $x \geq 0$, and zero otherwise. In this paper, we use the Generalized Canonical Ensemble (GCE) [17]:

$$f(U) = \beta_0 U + \frac{1}{2} \alpha (U - U_0)^2 \quad (6)$$

to generate points on $\beta_s(U)$. In thermodynamic stable regions, $\beta_s' < 0$. Thus, the distributions in the CE ($\alpha=0$) are Gaussian-like functions around E_f . However, in thermodynamic unstable regions such as two phase-coexistence regions, which cannot be sufficiently visited in the CE, we can use the GCE with $\alpha > \beta_s'(E)$ to stabilize the phase-coexistence regions, then reconstruct the data points of $\beta_s(U)$ in the regions.

The implementation of our algorithm is straightforward. Without losing generality, we choose U as the total potential energy of a physical system and define $T_s(U) = \beta_s^{-1}(U)$. The steps of our algorithm are as follows:

- (1) choose an initial T_0 (usually the highest temperature we desired), and then set $T_s(U) \equiv T_0$ and estimate $S(U) = U/T_0$;
- (2) simulate a segment of trajectory in the current $S(U)$ ensemble and generate an equilibrium sample;
- (3) choose a few $f_i(U)$ functions and calculate \bar{U}_i and fluctuations σ_i from the generated sample in the step 2

based on eq. (1).

(4) Interpolate $T_s(U)$ and linearly extrapolate it out of the current visited regions from the ending U_i from the obtained points $\{(\bar{U}_i, T_i)\}$ (and the derivative of $T_s(U)$, $\{T_i' = T_i^2 / \sigma_i^2\}$), then integrate $S(U)$ from the $T_s(U)$;

(5) Repeat the steps 2, 3, and 4 until the whole $T_s(U)$ and $S(U)$ are generated.

At the step 3, we may use CEs, for example, $f(U) = U/T_i$ with a series of $\{T_i\}$. Here, T_i is chosen from the current $T_s(U)$ function: the corresponding \tilde{U}_i , which satisfies the equation $T_i = T_s(\tilde{U}_i)$, must locate in the current visited regions and $|U_i - U_{i-1}| \sim \sigma_{i-1}$, where σ_{i-1} is the fluctuation of U at the temperature T_{i-1} . Also, we apply the mean and fluctuation at T_i when and only when the effective size of the weighted sample, B , is sufficiently large. Note that we can calculate the means and fluctuations at many T_i from the same sample without new simulations. The computational cost for the weighted mean is almost negligible compared with the generation of samples. In this step, we can check on whether larger α should be applied in applying the GCE to estimate \tilde{U}_f . For example, we first choose $\alpha=0$ to calculate (T_i, \tilde{U}_i) , then set $\beta_0 = T_i^{-1}$, $U_0 = \tilde{U}_i$, and use a positive α to calculate the means in the GCE ensemble, and finally compare the results to determine whether a larger α should be used to depress the possible deviation between the extreme point and the average U . The points (β_i, U_i) can be calculated using many samples generated in different segments of simulations in the step 3:

$$\bar{U} \leftarrow \frac{\sum \bar{U}^j B^j}{\sum B^j}, \quad (7)$$

where \bar{U}^j and B^j are the weighted mean of U and the corresponding B factor from the j th sample, respectively. The statistical error of \bar{U} , which is measured by the total $B = \sum B^j$ factor, can be suppressed by this way.

Before all regions are sufficiently visited, $T_s(U)$ in some regions may already converge if the B factor of the corresponding \bar{U} is large enough. For example, in normal physical systems, $T_s(U)$ at high temperatures can be obtained before the low-temperature regions are sufficient visited.

The necessary visiting time in the high- U regions, which is proportional to N even in the flat-histogram ensemble, can be further suppressed, so the simulation can focus its sampling more in the low- U region. We transform U to a new variable Y by $\frac{dY}{dU} = g_c\left(\frac{U - U_c}{b}\right)$, where the compressed function $g_c(z)$ slowly approaches zero when $z \rightarrow \infty$ and is almost unity when $z \leq 0$. Here, U_c and b are parameters to be determined. One possible selection is

$$g_c(z) = \frac{e^{-z} + 1}{e^{-z} + z + c}, \quad (8)$$

with the parameter $c \sim 2$ or 3 . The U distribution of Y -flat samples is $P(U) \sim 1/(U - U_c)$ when $U > U_c$. At the end of each segment of simulations, we reset U_c to be the lower boundary of the clear region and choose b in such a way that the expected visiting probability above U_c is comparable with that below U_c , and then use $\tilde{S}(Y)$ to generate the Y -flat sample. Here, $\tilde{S}(Y) = S(U) - \ln g_c\left(\frac{U - U_c}{b}\right)$. Although the

Y -flat simulations can still visit the high- U region ergodically, the majority of the simulation time is spent in the low- U region to refine $T_s(U)$. The compression transformation significantly decreases the required computational cost of our algorithm in very large systems.

3 Results

To verify the applicability of our algorithm, we applied it to a truncated-and-shifted Lennard Jones (LJ) fluid ($r_c = 2.5 \sigma_{ij}$) with the reduced density $\rho = 0.8$. U was selected as the total potential energy, and the DOS were calculated in the range between $T_l = 0.5$ and $T_h = 100$, corresponding to a very large range of U in large systems. The ensemble mean of U is regarded as exact if the error of \bar{U} is smaller than 5% of the fluctuation of U (i.e., $B > 400$). The Langevin thermostat was adopted to keep the system temperature constant and conformations were sampled every 10^5 MD steps. $T_s(U)$ was calculated and updated every 10^5 steps along with the parameters for the compressive transformation. In our simulations, the distance between neighboring U_i is much larger than in the WL-like methods [10] in high- T regions and large- N systems but automatically decreases in low- T regions and small N systems. The results are shown in Figure

1. The $T_s(U/N)$, $c_v \equiv \left(\frac{dT_s}{du}\right)^{-1}$, and $s(u) = S(u)/N$ quickly

converged in the large U regions, so the scaling relations were easily found. Here $u = U/N$. Figure 2 shows the efficiency of our approaches by the number of required MD steps $t \propto N^\gamma$ with $\gamma \approx 0.5$. Here, all ensemble means are required to reach the condition $B > 400$ so that their relative errors are smaller than 0.05.

As shown in Figure 3, at the lower temperature region, the obtained $T_s(U)$ is found to have multiple branches corresponding to the DOS of some separated parts in the conformational space. For a segment of simulated trajectory, we may know the branch of $T_s(U)$ the trajectory belongs to by calculating the corresponding $(T, \langle U \rangle_T)$ points. In our simulations, the trajectories generated in the right branch of $T_s(U)$ (liquid phase) have the expected flat histograms in the region $U > U_1$, but quickly transit into the low- U region (sol-

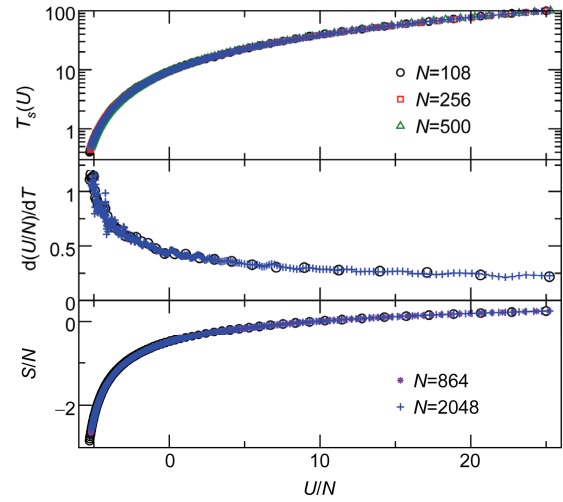


Figure 1 (Color online) The obtained T_s , $\left(\frac{dT_s}{du}\right)^{-1}$, and S are shown as functions of U in the LJ systems with different size N . The number of requiring points $\{(U_i, \beta_i)\}$ increases as N increases.

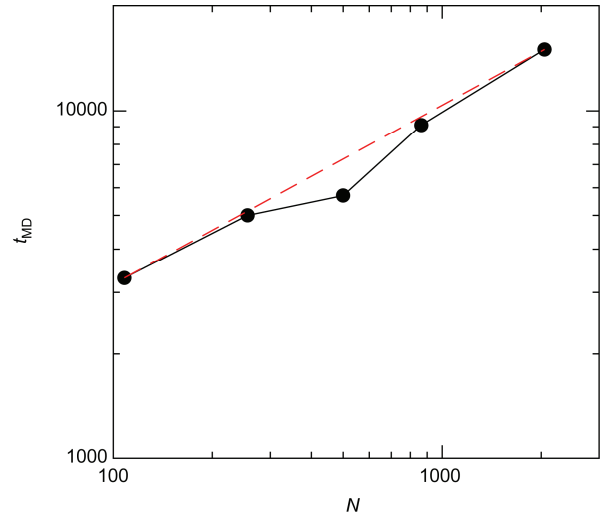


Figure 2 (Color online) The required MD steps for getting the same accuracy of $\beta_s(U)$ with the size of system N . The accuracy is measured by the B factors in estimating ensemble averages of U (see the text). The red dash line is from $t_{MD} \propto N^{1/2}$.

id phase). Similarly, simulations in the left branch of $T_s(U)$ (solid phase) homogeneously visit the $U < U_2$ region and transit into the liquid phase. The two branches have an overlap region, indicating that the potential energy of the liquid phase can be lower than that of the solid phase, and U definitely cannot serve as the order parameter for the liquid/solid phase transition. It is very hard to visit (and more difficult to identify) the low-energy liquid and high-energy solid conformations with the simulation techniques other than ours. Therefore, our approach, which is able to visit more parts of the configurational space and to identify them by forming their own DOS without a priorly defined order

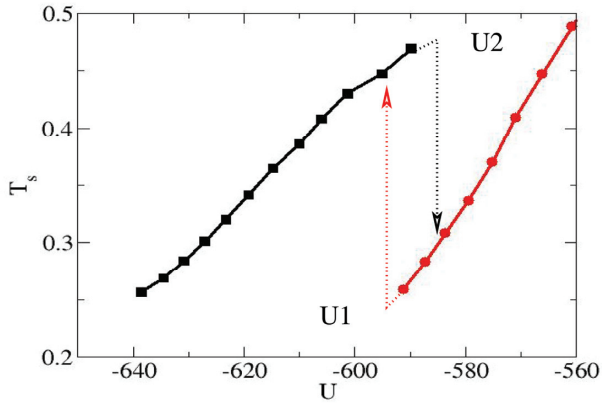


Figure 3 (Color online) $T_s(U)$ in the supersaturated liquid and the solid phases, respectively, of the LJ system with $N=108$. The transition can easily happen when $U < U_1$ (from liquid to solid) and $U > U_2$ (from solid to liquid).

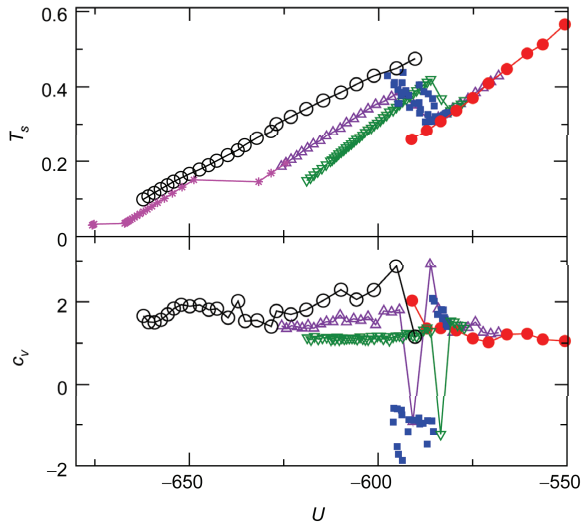


Figure 4 (Color online) The $T_s(U)$ function in a LJ system with $N=108$ and $\rho=0.8$ at the low-temperature region. Each branch of the function corresponds to a particular macroscopic phase. The blue-filled squares cover

the unstable liquid/solid coexistence region where $c_v \equiv \left(\frac{dT_s}{dU}\right)^{-1}$, which

can be thought as the heat capacity in the iso- U ensemble, are negative, as shown in the bottom panel. The green down-triangles correspond to the complete FCC solid, which is a metastable state in the current system. Some phase transitions from metastable solids to the stable solid (the black circles) can be found at lower temperatures. When $T_s \rightarrow 0$, more complex phase behaviors are expected.

parameter, provides a new approach for better understanding phase transition processes.

4 Discussion

It is possible to calculate the total $S(U)$ and $T_s(U)$ from the obtained $T_s(U)$ branches, $S(U) = \ln \sum_i \exp[S^i(U)]$, the

corresponding $T_s^{-1}(U) = \sum_i \omega_i(U) [T_s^i(U)]^{-1}$, where

$$S^i(U) = S_0^i + \int_{U_{i0}}^U T_s^i(E) dE \quad (9)$$

and $T_s^i(U)$ are the i th branch of the entropy and temperature, respectively. Here $\omega_i(U) = \exp[S^i(U)] / \sum_i \exp[S^i(U)]$. In the calculation, the integration constants $\{S_0^i\}$ must be determined from another simulation technique. However, it should be noted that even in the exact $S(U)$ ensemble, free transition between states still cannot be guaranteed, because the application of $S(U)$ does not change $P(r)$ inside each iso- U subspace. This is a general limitation in all low-dimension flat-histogram ensemble methods. The algorithms with good stabilities (e.g., satisfying the detailed balance in a long time) are helpful for analyzing and overcoming the limitation. For example, our approach generates $T_s(U)$ and $S(U)$ (with an arbitrary constant) in each visited subspace instead of the expected total $T_s(U)$ and $S(U)$ for the whole conformational space. Therefore, although the equilibrium among subspaces is difficult to reach, it still provides many useful results and might construct the total $S(U)$ by determining the relative constants of $S(U)$ for each subspace. Also, we may use the generalized ensembles described by eq. (6) to find more metastable states. For example, besides the stable liquid and solid phases in the LJ system, we detected some metastable solid states with different crystalline structures as well as the unstable multiphase coexistence, which are shown in Figure 4. These complex solid phases appear because the system is fixed at the density $\rho=0.8$, so the interparticle distance in the Face-Center-Cubic (FCC) structure does not match the equilibrium distance of the LJ potential, $r_0=2^{1/6}$.

The flat-histogram ensembles based on multiple collective variables are expected to better enhance the sampling than the ensembles based on a single variable [11,19]. The computation complexity of our algorithm is $O(N^{\gamma d})$, where d is the number of the collective variables and $\gamma=1/2$, which is much faster than the WL-like with $\gamma=1.5$ to 3. Recently, Vogel et al. [20] combined the original WL algorithm with the Replica Exchange Method (REM) to form a parallel WL sampling method, which makes application of the WL algorithm in large systems much efficient. It is also possible to combine the FAFE algorithm with the REM to further improve the efficiency of sampling. We may apply replicas in different U ranges [20], or in different branches of $T_s(U)$, then exchange conformations among these replicas. It might provide a possibility to overcome the main limitation in the application of REM, which requires $O(N^{1/2})$ replicas to achieve sufficient exchange acceptance ratio [21].

This work was supported by the National Natural Science Foundation of China (Grant No. 11175250) and by the Open Project Grant from the State

Key Laboratory of Theoretical Physics. Zhou X thanks the financial support of the Hundred of Talents Program in Chinese Academy of Sciences.

- 1 Fang Y. A Gibbs free energy formula for protein folding derived from quantum statistics. *Sci China-Phys Mech Astron*, 2014, 57: 1547–1551
- 2 Luo L F. Quantum theory on protein folding. *Sci China-Phys Mech Astron*, 2014, 57: 458–468
- 3 Gong P, Wang X, Shao Y, et al. Ti-Zr-Be-Fe quaternary bulk metallic glasses designed by Fe alloying. 2013, 56: 2090–2097
- 4 Berg B, Neuhaus T. Multicanonical –algorithms for first order phase transitions. *Phys Lett B*, 1991, 267: 249253
- 5 Wang F, Landau D. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys Rev Lett*, 2001, 86: 2050–2053
- 6 Laio A, Parrinello M. Escaping free-energy minima. *Proc Natl Acad Sci USA*, 2002, 99: 12562–12566
- 7 Yan Q, de Pablo J. Fast calculation of the density of states of a fluid by Monte Carlo simulations. *Phys Rev Lett*, 2003, 90: 035701
- 8 Zhou C, Bhatt R N. Understanding and improving the Wang-Landau algorithm. *Phys Rev E*, 2005, 72: 025701
- 9 Zhou X, Jiang Y, Kremer K, et al. Hyperdynamics for entropic systems: Time-space compression and pair correlation function approximation. *Phys Rev E*, 2006, 74: 035701
- 10 Kim J, Straub J E, Keyes T. Statistical-temperature Monte Carlo and molecular dynamics algorithms. *Phys Rev Lett*, 2006, 97: 050601
- 11 Zhou C, Schulthess T C, Torbrügge S, et al. Wang-landau algorithm for continuous models and joint density of states. *Phys Rev Lett*, 2006, 96: 120201
- 12 Jeżewski W. Stationary and nonstationary properties of evolving networks with preferential linkage. *Phys Rev E*, 2002, 66: 067102
- 13 Gao Y Q. An integrate-over-temperature approach for enhanced sampling. *J Chem Phys*, 2008, 128: 064105
- 14 Zhang C, Ma J. Enhanced sampling in generalized ensemble with large gap of sampling parameter: Case study in temperature space random walk. *J Chem Phys*, 2009, 130: 194112
- 15 Gong L, Zhou X. Structuring and sampling complex conformation space: Weighted ensemble dynamics simulations. *Phys Rev E*, 2009, 80: 026707
- 16 Gong L, Zhou X. Kinetic transition network based on trajectory mapping. *J Phys Chem B*, 2010, 114: 10266–10276
- 17 Xu S, Zhou X, OuYang Z C. Parallel tempering simulation on generalized canonical ensemble. *Commun Comput Phys*, 2012, 12: 1293–1306
- 18 Junghans C, Perez D, Vogel T. Molecular dynamics in the multicanonical ensemble: Equivalence of Wang-Landau sampling, statistical temperature molecular dynamics, and metadynamics. *J Chem Theory Comput*, 2014, 10: 1843–1847
- 19 Mastny E A, de Pablo J J. Direct calculation of solid-liquid equilibria from density-of-states Monte Carlo simulations. *J Chem Phys*, 2005, 122: 124109
- 20 Vogel T, Li Y W, Wüst T, et al. Generic, hierarchical framework for massively parallel Wang-Landau sampling. *Phys Rev Lett*, 2013, 110: 210603
- 21 Earl D J, Deem M W. Parallel tempering: Theory, applications, and new perspectives. *Phys Chem Chem Phys*, 2005, 7: 3910–3916